

Information Retrieval in Biomedicine: Natural Language Processing for Knowledge Integration. Edited by Violaine Prince and Mathieu Roche. Hershey, NY: Medical Information Science Reference; 2009. 432 p. \$225.00. Perpetual access: \$340.00. Print + perpetual access: \$450.00. ISBN: 978-1-60566-274-9.

For a number of years, librarians have heard that natural language processing (NLP) will revolutionize information management and retrieval in health care settings. The goal of the two editors, professors at the University Montpellier 2 in France, is to compile research that librarians and health information systems administrators and developers will find useful in incorporating data management systems solutions into their organizations. This book provides relevant theoretical frameworks and empirical research findings in NLP according to linguistic granularity and presents original applications.

Both editors demonstrate expertise in the field of computer science. Prince headed the French National University Council for Computer Science and leads the NLP research team at the Laboratoire d'Informatique, de Robotique et de Microelectronique de Montpellier. She specializes in NLP and cognitive science. Roche's research interests include text mining, information retrieval (IR), terminology, and NLP for schema mapping.

NLP is a subfield of computer science that addresses the operation and management of texts, as input or output of computational devices. The scientific community is interested in NLP for the following uses: IR and knowledge extraction, knowledge integration to existing devices, and use and application of existing knowledge structures for IR services. Scientific literature encompasses so much knowledge that only computer-based systems can browse and filter it. A regular search engine does not have the complexity to undertake complex scientific queries. Knowledge classifications, taxonomies with hierarchical ties

between knowledge items, provide ontologies at a high cost in human labor and involvement. NLP tools reorganize artificial intelligence (AI) techniques to focus on linguistic-conceptual relationships, rather than primarily textual analysis. Knowledge integration translates synonymous terms using data and text mining to complete or correct existing knowledge structures. Medical literature contains a substantial number of acronyms. In NLP, unlike mathematical originated formalisms, a concept can be addressed through a variety of words and phrases that are not exactly equivalent. Retrieving the relevant set of texts from a complex query must not only rely on words, but also grasp ideas expressed by distinct strings of words or phrases and necessitates topical classification. More than half of the scientific literature is written in non-English languages. Key BioNLP domain resources include PubMed, ontologies and thesauri (e.g., GeneOntology), and the Medical Subject Headings (MeSH) Thesaurus.

The extraction process for IR and knowledge management (KM) is the same. However, IR results in raw data, and KM results in machine operable data. The two properties that define this book are emphasizing NLP as the main methodological issue and studying the interaction between NLP and its application domains of science and medicine. NLP goes far beyond the concept of word computing to include sentence level and discourse, segment, or text level. Sentence meaning results from word interactions as well as word meanings. Paragraph positions and grammar demonstrate the intentions of the human authors and are generally ignored by most computational techniques. NLP theories provide an attempt to address how text organization shows the dependence of language on its nonlinguistic environment. Such theories include discourse relations theory, discourse rhetorical structures theory, and speech acts theory. BioNLP has developed its own characteristics to process the domain language of biology and medicine.

The target audience includes graduate school professors and students, NLP researchers, AI researchers, terminologists, linguists, health information systems specialists, and the BioNLP community. Librarians in academic settings serving the aforementioned may find this book useful in their collections, particularly if they serve those interested in European research, because the majority of the twenty-two chapter authors are international.

The book is divided into five sections: "Works at a Lexical Level," "Crossroads between NLP and Ontological Knowledge Management"; "Going Beyond Words and NLP Approaches Involving the Sentence Level"; "Pragmatics, Discourse Structures and Segment

Level as the Last Stage in the NLP Offer to Biomedicine"; "NLP Software for IR in Biomedicine"; and "Conclusion and Perspectives." The book begins with a chapter on text mining for biomedicine that sets the stage. Following the preface and chapters are sections on compilation of resources, about the contributors, and the index. References are predominantly from the last ten years, with some as new as 2008, though most are older. Some are needlessly cited twice. Contributors represent ongoing research from across nations and disciplines. The index does not include cross-references, which is odd considering the topic of the book. There are some typos or misspellings dispersed throughout the text. Graphs and illustrations

are relevant and enhance the reader's understanding of concepts. The numerous mathematical formulas will appeal to the target audience.

Prince and Roche have succeeded in compiling an original work in the field of BioNLP. There appears to be no advantage to the purchase of perpetual access, although potential certainly exists for an enhanced electronic version.

*Martha F. Earl, MSLS, AHIP,
mearl@utmck.edu,
Preston Medical Library, University
of Tennessee Graduate School of
Medicine, Knoxville, TN*

DOI: 10.3163/1536-5050.98.2.020



R2: Converting Librarians to Cybrarians

The R2 Digital Library from Rittenhouse makes it easy for librarians everywhere to offer more than 1,000 health sciences titles online.

SAVE MONEY. Avoid "all-or-nothing" prices. Add more titles immediately with the push of a button.
SAVE TIME. Our easy-to-use interface meets needs of librarians. If you can surf the web, you can use R2.
MANAGE SMARTLY. Monitor usage statistic in real time, so you can make informed decisions.
 Move into library management's digital age with a name you can trust: Rittenhouse.

Rittenhouse

BOOK DISTRIBUTORS

DISTRIBUTING THE LEADING SCIENTIFIC, TECHNICAL AND MEDICAL PUBLISHERS

511 Feheley Dr., King of Prussia, PA 19406 • 800-345-6425 • www.r2library.com • www.rittenhouse.com